

Distance Sampling with a Random Scale Detection Function

Appendix S1: Comparison between two likelihood formulations for the random scale detection function

Cornelia S. Oedekoven Jeffrey L. Laake Hans J. Skaug

7 A scale mixture model likelihood for the pdf

A seemingly different approach compared to the detection function likelihood L_g is to apply the scale mixture to the probability density eq (1) rather than to the detection function. This fits directly into the classical line transect theory where animal density is calculated from eq (3). The motivation for this scale mixture is to obtain a more flexible parametric family for the pdf $f(x)$. In the following section we will show that – in the case of the half-normal detection function in combination with normal random effects – this is equivalent, in the sense of producing identical estimates of animal density, to that of the method outline in Section 3 of the main manuscript.

The individual level pdf, in absence of truncation ($w = \infty$), corresponding to eq (5) is

$$f(x|\epsilon) = 2 \left(\sqrt{2\pi}\gamma(\epsilon) \right)^{-1} \exp \left(-x^2/(2\gamma(\epsilon)^2) \right), \quad x \geq 0.$$

Applying the scale mixture of Section 3 yields the pdf

$$f(x) = \int_{-\infty}^{\infty} f(x|\epsilon) N(\epsilon, 0, \sigma_{\epsilon}) d\epsilon,$$

which taking data truncation into account yields the likelihood

$$L_f(\beta, \sigma_{\epsilon}) = \prod_{i=1}^n \frac{\int_{-\infty}^{\infty} f(x_i|\epsilon) N(\epsilon, 0, \sigma_{\epsilon}) d\epsilon}{\int_{-\infty}^{\infty} \int_0^w f(u|\epsilon) du N(\epsilon, 0, \sigma_{\epsilon}) d\epsilon}. \quad (13)$$

We now explain why and how L_g and L_f differ. For this purpose we need to distinguish between the 'detected' population and the (full) population which consists of all individuals. Both these populations are

considered to be of infinite size, with the detected population resulting from having run the survey for a very long time period. The detection process acts as a 'filter' which affects the distribution of both x and ϵ . For x this is the well known fact that if x is uniformly distributed in the full population, it is distributed according to eq (1) in the detected population. Similarly, is proved in Section 9 below that if $\epsilon \sim N(0, \sigma_\epsilon)$ in the full population, then $\epsilon \sim N(\sigma_\epsilon^2, \sigma_\epsilon)$ in the detected population. This implies a shift of $\gamma(\epsilon) = \exp(\beta + \epsilon)$ towards larger values of the scale parameter γ , reflecting the fact that individuals with a large value of γ are more likely to be detected. The likelihood L_f , on the other hand, is based on the directly contradicting assumption that $\epsilon \sim N(0, \sigma_\epsilon)$ in the detected population. However, since the only difference between L_g and L_f is a right-shift in the distribution of ϵ in the detected population it can be proved (see Section 9 below) that they give identical estimates of σ_ϵ , and the estimates of the intercept β are related through

$$\beta_f = \beta_g + \sigma_\epsilon^2, \quad (14)$$

where β_f and β_g are the intercepts obtained from L_f and L_g , respectively.

8 Comparing the f and g formulations via simulation

With simulation we compare the results from the R and ADMB code obtained with the two different likelihood formulations, L_g eq (6) and L_f eq (13). For this purpose, we simulated $n=500$ distances from a half-normal detection function with random scale ($\beta_g = 2; \log(\sigma_\epsilon) = -0.5$) and without any truncation (i.e. $w = \infty$). We then fit the model to these data using eq (6) solely with R code and then with the ADMB implementations using likelihood eqs (6) and (13). Fig. 4 shows the plot for each fitted model; Table 3 demonstrates that L_g and L_f produce identical results with the ADMB code and the R and ADMB code produce nearly identical results for eq (6).

9 Connection between the f and g formulations

The result referred to above in eq (14) applies to the situation where the detection function is of the form $g(x|\gamma) = k(x/\gamma)$, where $k()$ is some base detection function and $\gamma > 0$ is a scale parameter. Under this model the perpendicular density for line transects is $f(x|\gamma) = \{\int_0^\infty g(x|\gamma) dx\}^{-1} g(x|\gamma) = \gamma^{-1} \{\int_0^\infty k(u) du\}^{-1} k(x/\gamma)$. The second requirement for the result to hold is that the scale parameter γ has a log-normal distribution given by eq (9). For simplicity we ignore the regression part in eq (9) such that $\gamma(\epsilon) = \exp(\epsilon)$, where $\epsilon \sim N(\beta, \sigma_\epsilon)$.

Consider two sampling mechanisms:

1. Each individual in the population (whether detected or not) has detection function $g\{x|\gamma(\epsilon)\}$, where $\epsilon \sim N(\beta, \sigma_\epsilon)$. This gives rise to the likelihood in eq (6).

2. Each individual *among the detected individuals* has perpendicular density $f\{x|\gamma(\epsilon)\}$, where $\epsilon \sim N(\beta, \sigma_\epsilon)$. This gives rise to the likelihood in eq (13).

The difference between the two is the subgroup to which the assumption about lognormality is applied. The following result shows that the likelihoods are linked parametrically.

Theorem Let $p_g(x; \beta, \sigma_\epsilon)$ and $p_f(x; \beta, \sigma_\epsilon)$ refer to the density of perpendicular distances truncated to the interval $[0, w]$ under sampling mechanisms 1) and 2) above, respectively. We then have $p_f(x; \beta, \sigma_\epsilon) = p_g(x; \beta - \sigma_\epsilon^2, \sigma_\epsilon)$, and $\beta_f = \beta_g + \sigma_\epsilon^2$.

Proof We put $\gamma(\epsilon) = \exp(\epsilon)$, and get

$$\begin{aligned} p_g(x; \beta - \sigma_\epsilon^2, \sigma_\epsilon^2) &= \frac{\int g(x|\gamma(\epsilon)) N(\epsilon, \beta - \sigma_\epsilon^2, \sigma_\epsilon) d\epsilon}{\int \int_0^w g(u|\gamma(\epsilon)) du N(\epsilon, \beta - \sigma_\epsilon^2, \sigma_\epsilon) d\epsilon} \\ &= \frac{\int \gamma(\epsilon)^{-1} g(x|\gamma(\epsilon)) \exp(\epsilon) N(\epsilon, \beta - \sigma_\epsilon^2, \sigma_\epsilon) d\epsilon}{\int \int_0^w \gamma(\epsilon)^{-1} g(u|\gamma(\epsilon)) du \exp(\epsilon) N(\epsilon, \beta - \sigma_\epsilon^2, \sigma_\epsilon) d\epsilon} \\ &= \frac{\int \gamma(\epsilon)^{-1} g(x|\gamma(\epsilon)) N(\epsilon, \beta, \sigma_\epsilon) d\epsilon \exp(\beta - \frac{1}{2}\sigma_\epsilon^2)}{\int \int_0^w \gamma(\epsilon)^{-1} g(u|\gamma(\epsilon)) du N(\epsilon, \beta, \sigma_\epsilon) d\epsilon \exp(\beta - \frac{1}{2}\sigma_\epsilon^2)} \\ &= p_f(x, \beta, \sigma_\epsilon) \end{aligned}$$

10 Discussion

We have illustrated two different likelihood functions, L_g and L_f . With the former, the random effects distribution is associated with the population of all animals, while with L_f the random effects distribution refers only to detected animals. In the epidemiological literature this distinction has been noticed (Glidden and Liang, 2002) in the context of what is known as “ascertainment bias”. Our relationship (14) quantifies the ascertainment bias in our situation. By highlighting the possibility of using different likelihood functions we hope we will make the reader, and potential user of these methods, aware that it is necessary to keep in mind which population the random effects distribution refers to.

References

- Fournier, D. A., S. H. J., Ancheta, J., Ianneli, J., Magnusson, A., Maunder, M. N., Nielsen, A., and Sibert, J. (2012). Ad model builder: using automatic differentiation for statistical inference of highly parameterized complex nonlinear models. *Optimization Methods and Software* **27**, 233–249.
- Glidden, D. V. and Liang, K.-Y. (2002). Ascertainment adjustment in complex diseases. *Genetic epidemiology* **23**, 201–208.

Table 3: Comparison of random scale model fitting results from R and ADMB using simulated data from a random scale half-normal with $\beta_g = 2$ and $\log(\sigma_\epsilon) = -0.5$.

Code	Likelihood eq	Log-likelihood	β_f	β_g	$\log(\sigma_\epsilon)$
R	eq (6)	-1614.82	NA	1.950	-0.538
ADMB	eq (6)	-1614.79	NA	1.948	-0.536
ADMB	eq 13	-1614.79	2.290	1.948	-0.536

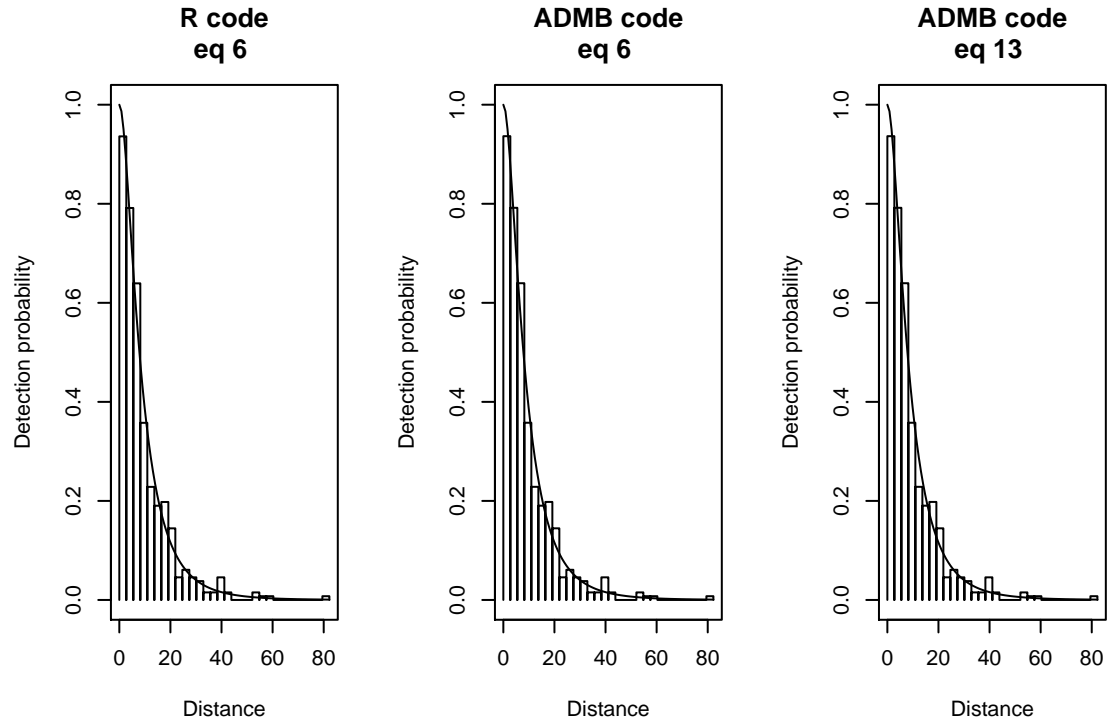


Figure 4: Random scale detection functions fitted to simulated data with R and ADMB code using the likelihood eq (6) and with the ADMB code using eq (13).